

UNDERSTANDING SPOKEN DISCOURSE – MS 4296

F. Cornish, CNRS ERSS, UMR 5610 & Department of Language Sciences, University of Toulouse-Le Mirail, 5, Allées Antonio Machado, 31058 Toulouse Cedex 09, France

ABSTRACT

Spoken discourse is the ongoing, situated interpretation of a speaker's communicative intentions, of which the addressee's expected and actual reactions are an integral part. The creation of discourse is thus a joint endeavor, involving the active cooperation of all the participants. The textual record on which this constructive activity is based, in conjunction with the invocation of a relevant context, is constituted not only by the verbal content of the utterances produced, but also by non-verbal signals. According to Clark (1996), the text flow is divided between two simultaneously operative tracks: a primary "official business" track, and a secondary discourse-management track. Evidence for the existence and specialization of these two tracks is given via extracts from two radio broadcast discussions.

1. Introduction: the contextual specificity of spoken interaction

A characterization of spoken discourse requires a specification of the very different contexts of utterance which obtain in the production and reception of spoken and written discourse. These configurations largely motivate the rather different properties of spoken and written discourse, even allowing for equivalence of register and formality: see the well-documented account of cross-language spoken syntax and discourse in Miller & Weinert (1998).

Speech prototypically involves face-to-face interaction between two or more participants who share a spatio-temporal environment. This, together with a common cultural and personal background in the case of conversationalists who know each other well, provides a rich contextual common ground allowing the speaker to avoid having to verbalize a number of aspects of his or her message. Concomitantly, this common ground enables the discourse participants to rely to a large extent on non-verbal signaling, in tandem with and even, on occasion, in place of, the verbal textualisation of a given utterance. Planning time, as well as "understanding" time, is naturally minimal and at a premium – and a great many features of spontaneous speech flow from this key factor. Moreover, both speech and writing are normally designed by the user so as to be readily understood by the addressee (cf. the notion of "recipient design"). Indeed, according to Clark (1996) and other linguists, conversation and communication in general is a fundamentally joint activity, involving the active participation of the interlocutors and the coordination of their actions (verbal as well as non-verbal).

What I have just (very briefly) characterized is of course the prototypical instance of spoken interaction. There are obviously other less prototypical types of spoken discourse: for example, speaking on the telephone, where the participants share a time frame (adjusting for time zone differences when the call is international), but not a spatial one, where only two participants are involved, and where the communication is 'ear-to-ear' rather than face-to-face (no non-vocal gestures or visual percepts are possible): see Drummond & Hopper (1991) for a discussion of miscommunication over the telephone; and speaking in a formal situation (a speech, lecture and so forth) in front of a group of people in circumstances where convention does not normally allow for verbal exchange and interaction.

In written discourse, on the other hand, there is by definition no common spatio-temporal ground between the writer and their reader(s). Since this is the case, and since inevitably there will be little or no opportunity to use non-verbal signals, the text used will need to be relatively explicit - since the textual input is confined to the verbal content, in conjunction with punctuation and various graphic devices. The much greater availability, in principle, of planning time allows the writer to review and to amend their written production.

2. The context of spoken discourse, the distinction between *text* and *discourse*, and their roles in understanding

It is useful in analyzing spoken (as well as written) discourse understanding to draw a three-way distinction between the dimensions of *text*, *discourse* and *context*. Definitions which I find helpful are given under (1) below (see Cornish, 1999: §2.3 and 2003: §2 for further development and illustration of the ‘text’/‘discourse’ distinction, and its importance for anaphora; also Edmondson, 1981: 4, Seidlhofer & Widdowson, 1999, and Werth, 1999: ch. 5 in connection with the notion ‘context’, on ‘Common Ground’ within his ‘Text Worlds’ framework):

(1) *Text, discourse and context*

Text: the connected sequence of verbal signs and non-verbal signals in terms of which discourse is co-constructed by the participants in the act of communication.

Discourse: the hierarchically structured, situated sequence of indexical, propositional, utterance and illocutionary acts carried out in pursuance of some communicative goal, as integrated within a given context.

The *context* is subject to an ongoing process of construction and revision as the discourse unfolds. It is through the invocation of a relevant context (which is partly determined by the nature of the co-text at issue, as well as by its genre) that the hearer or reader is able to convert the connected sequence of textual cues that is *text* into *discourse*. (Extract (slightly amended) from Cornish, 2003:3).

The *text* is the perceptible record of at least one utterance act (whether realized in terms of a verbal, linguistic trace or of a non-verbal trace - which may be gestural, sensory-perceptual or prosodic). See especially Clark’s (1996) chapter 6 on non-verbal signals and their different kinds of functions in discourse. Clark draws a highly relevant distinction between two simultaneously functioning textual “tracks” which operate in spoken discourse: a primary track, where the “official business” of the transaction at hand is being conducted; and a secondary “meta-discursive” or discourse management track, where participants make explicit the purposes and functions of the preceding and ongoing talk. The functions of either track may be realized via verbal and non-verbal signals (whether in tandem or individually). We shall be looking at examples of this dual-track structure in operation shortly.

The notion of *text* is close to what Gumperz (1992: 234) calls “contextualization cues”. The discourse partners make use of this record (a dual-track one, according to Clark, 1996), in conjunction with their invocation of a relevant *context* in cognitive terms, in order to create *discourse*.

Discourse, on the other hand, refers to the hierarchically-structured, mentally-represented product of the sequences of utterance, propositional, illocutionary and indexical acts which the participants are carrying out as the communication takes place. Such sequences have as their *raison d'être* the accomplishment of some particular overall communicative goal (see Parisi & Castelfranchi 1977).

The crucial point about this distinction is that discourse is a (re-)constructive, and therefore highly probabilistic enterprise: from the addressee's perspective, it is by no means a question of simply directly **decoding** the text in order to arrive at the fully-fledged message originally intended by the addressor. Indeed, the addressee actively contributes both to the text and to the discourse via their phatic signals, indications of (mis)understanding, and other reactions to the speaker's moves. 'Meaning' does not lie "in" the text, it has to be constructed by the addressee (and the speaker!) via the text and an appropriate context (cf. Coupland et al., 1991: 5). In any case, the text is often, if not always, both incomplete and indeterminate in relation to the discourse which may be derived from it in conjunction with a context.

3. Some aspects of understanding spoken discourse

3.1 Inferring propositional content and illocutionary force

In what follows, I shall be examining instances of (mis)understanding which occur, and are manifest, **within** conversations. Thus it is the discourse participants themselves whose monitoring of the discourse being co-constructed is at issue here. I am adopting the principle that it is when misunderstandings, disagreements or disruptions generally are manifest in the textual record of a conversation that the way in which discourse normally operates may be seen most clearly (cf. Coupland et al., 1991). Let us analyze an initial occurrence of such a phenomenon. Here is a segment from the BBC Radio 4 cultural discussion program *Start the Week*. The previous speaker (Caroline Quinn) has been arguing that the alienation of Black people in the United States is not due to a single factor, but has a variety of causes; that the situation is improving for all racial groups in the US, and that differences in degrees of integration into American society are in part due to differences in the "cultural inheritance" which each group brings with it. Homi Barber is then given the floor (for a second time) by the presenter, Melvin Bragg. Notational symbols used here are as follows: '-': pause; '- -': double pause; upper case letters: strongly accented syllable; '[...]': simultaneous speech; '=': latching; '(a)': elision of "a"; '.hhh' = sharp intake of breath. See Cameron's (2001: 31-44) chapter 3 and Schiffrin's (1994: 422-438) Appendix 2 for details of spoken discourse transcription.

- (2) HB: Kate Kate – Caroline – you know I'm SURE you didn't mean it but sometimes – cultural inheritance shades off into biological inheritance – in in the States you know people say .hhh – Blacks are in some – inherent way – inferior – an' there's a lot of – a lot of a lot of stuff going around now of course – American Blacks came as slaves it's not what they brought with them it was what they were not – Able to bring with them they were [snatched – no but - - -]
- CQ: [it's no – it's nobody's fault] =
- HB: = no I'm not saying it's anybody's fault – but I'm just saying you know that that's the brute - historical - FACT...

In track 1 here, Caroline Quinn asserts that the fact that Black slaves were forced to sail to America without taking any of their possessions with them was not **their** fault (this is the proposition actually intended to be conveyed by the speaker here). It is no doubt the extreme sensitivity of the issue (racism towards Blacks) that has motivated the use of the indefinite negative with wide scope, *nobody*, here. (This is also the motivation for HB's intake of breath in line 3, just before his presentation of the (racist) view of Blacks in America). And in track 2, the same speaker is rejecting what she sees as Homi Barber's illocutionary stance in the extract – justifying and even seeming to condone the fact that Black people in the US are still not considered as “real” Americans. Barber's response (in track 2) is to assert that this is not **his** view (notice his repetition of the verb *say* in “I'm (not) saying...”), but that the source of the way Black Americans are viewed today is an objective, historical fact; and in doing so, he is clearly rejecting Caroline Quinn's implied interpretation of his illocutionary stance in his first turn.

3.2 *Inferring intended interactional moves and acts*

Occasionally, speakers and their addressee(s) disengage themselves from the discourse which they are creating, to establish meta-discursively what the relation is between their adjacent moves. This occurs in the “second” track, then, in Clark's (1996) terms. When this happens, we get an explicit view of how the discourse participants are interpreting each other's utterances (see also Cameron, 2001: 116). An example occurred during a discussion about “the Devil” in another edition of the BBC Radio 4 program *Start the Week* (22 April 1996). Here the presenter, Melvin Bragg (MB), is picking up on Peter Stanford's characterization of the use of “the Devil” by the mediaeval Church as a means of social control:

- (3) MB: yes – the interesting thing is that was it already coercive an' repressive – or did it actually call out something that's in people anyway – I mean the way you put it is that this is an authority this is a Church behaving as the Church behaved in many many different ways – in great ways an' in wicked ways – but this is a Church behaving in a very authoritarian way – in saying “you will follow us – or we will er – we will – get you and we've got the man to get you – he's called “the Devil”” – but – isn't there something else there – isn't it a recognition of what's part of human nature and it was a BRILLIANT metaphor – just as Christianity is full of brilliant metaphors as to what human nature is about so – .hhh it's more positive in a way than [what you're saying]
- PS: [well – well] it wouldn't have worked would it – un un unless they were actually tapping into something that people wanted to believe – and if you think about it =
- MB: = so you're agreeing =
- PS: = I'm agreeing – but if you think about the – concept of evil...

Melvin Bragg, in his first turn, is objecting to what he sees as a one-sided view of the status of the Devil in the mediaeval Church, according to Peter Stanford's earlier characterization: he is espousing the view that, rather than “imposing” the Devil on believers in an authoritarian manner, the mediaeval Church more intelligently used the Devil as a metaphor for something that it recognized as already being part of human nature. Given that it is the former view which Melvin Bragg understands Peter Stanford to be adopting, he clearly expects him to disagree with his objection. Yet Stanford's response is in conformity with this “objection”, so it cannot count as a “disagreement”. As soon as Bragg realizes this, he interrupts Stanford's response and asks him (via a

declarative request) for confirmation of this interpretation. Stanford then immediately gives it by repeating the “agreement” statement, then moves straight on to a development of the main point he had begun making at the point of interruption. This is achieved via a repetition of the actual words of his preceding final utterance (with the conjunction *and* replaced by the adversative *but*): ... *and if you think about it...* This is clear evidence of the “two-track” structure of textual development, as argued by Clark (1996), since the final conjunct of the pre-interrupted segment by Peter Stanford is repeated virtually verbatim immediately after the interruption. This indicates that the primary, “official business” track has not in fact been interrupted, but that it is the secondary, meta-discursive one which contained the interruption.

4. Conclusion

As Clark (1996) in particular emphasizes throughout his book, discourse is a joint endeavor, not the individual responsibility of the speaker, where the addressee has a merely passive role in decoding his/her utterances. The textual record (verbal content of the utterances as well as meaningful gestures, prosody, phatic and other vocalizations) radically underdetermines the discourse which the participants are jointly constructing as the text unfolds in a particular context. The discourse constructed at any given point in this unfolding is a tentative, probabilistic affair, and is not only subject to continual modification in terms of updating as each new utterance is encountered; but the immediately preceding discourse at a given point may also be revised and re-constructed retroactively, as a function of a subsequent discourse act or move. This may occur when a participant, encountering another’s reaction to what s/he is attempting to say, realizes the latter has misunderstood their propositional content, illocutionary stance or the nature of their move (see the examples under (2) and (3) above). In triggering this process of updating, revision and (re-)negotiation, discourse particles, vocal and visual gestures, pausing and prosody generally, all assume a crucial significance. The dual-track structure of textualisation postulated by Clark (1996) makes possible this parallel management of the discharge of “official discourse business” (track 1), on the one hand, and the signaling of discourse organization (track 2), on the other.

References

- Cameron, D. (2001). *Working with spoken discourse*. London & New York: Sage Publications.
- Clark, H.H. (1996). *Using language*. Cambridge: Cambridge University Press.
- Cornish, F. (1999). *Anaphora, discourse and understanding. Evidence from English and French*. Oxford: Clarendon Press.
- Cornish, F. (2003). ‘The roles of (written) text and anaphor-type distribution in the construction of discourse’. *Text* 23(1), 1-26.
- Coupland, N., Wiemann, J.M. & Giles, H. (1991). ‘Talk as “problem” and communication as “miscommunication”; an integrative analysis’. Ch. 1 in Coupland, N., Giles, H. & Wiemann, J.M. (eds.) *“Miscommunication” and Problematic Talk*. London & New Delhi: Sage Publications, 1-17.
- Drummond, K. & Hopper, R. (1991). ‘Misunderstanding and its remedies: telephone miscommunication’. Ch. 15 in Coupland, N., Giles, H. & Wiemann, J.M. (eds.)

- “Miscommunication” and Problematic Talk. London & New Delhi: Sage Publications, 301-314.
- Edmondson, W. (1981). *Spoken discourse. A model for analysis*. Harlow & New York: Longman Ltd.
- Gumperz, J. J. (1992). ‘Contextualization and understanding’. Ch. 8 in Duranti, A. & Goodwin, C. (eds), *Rethinking context. Language as an interactive phenomenon*, Cambridge: Cambridge University Press, 229-252.
- Miller, J. & Weinert, R. (1998). *Spontaneous spoken language. Syntax and discourse*. Oxford: Clarendon Press.
- Parisi, D. & Castelfranchi, C. (1977). ‘The discourse as a hierarchy of goals’. *Signs of Change* 1(2), 31-67.
- Schiffrin, D. (1994). *Approaches to discourse*. Oxford: Blackwell.
- Seidlhofer, B. & Widdowson, H. (1999). ‘Coherence in summary: the contents of appropriate discourse’. In Bublitz, W., Lenk, U. & Ventola, E. (eds.) *Coherence in spoken and written discourse*. Amsterdam & Philadelphia: John Benjamins, 205-219.
- Werth, P. (1999). *Text worlds: Representing conceptual space in discourse*. Harlow & New York: Pearson Education Ltd.

Keywords: Spoken discourse; Understanding; Text; Context; Coherence; Non-verbal signals; Interaction; Coordination; Textualisation; Propositional content; Illocutionary force/stance; Discourse management; Primary textual track; Secondary textual track.

Francis Cornish: brief biography

Francis Cornish gained a Ph.D in Linguistics in 1982 from the University of Sussex, and in 1997, a Postdoctoral Degree in the Language Sciences (‘Habilitation à Diriger des Recherches en Sciences du Langage’) from the University of Toulouse-Le Mirail, France.

He is the author of two books on anaphora in a discourse context (*Anaphoric Relations in English and French: a discourse perspective*, London & Canberra: Croom Helm 1986, and *Anaphora, Discourse, and Understanding. Evidence from English and French*, Oxford: Clarendon Press (Oxford University Press) 1999) and has written numerous articles on anaphora, deixis, reading, agreement, text and discourse structure, as well as information structure, reference, predication and the lexical semantics-syntax interface. He is currently working on the topic of null complements, and their effects on the Aktionsart of the host predicate’s containing predication. He works within the framework of the developing model Functional Discourse Grammar, an offshoot of standard Dikian Functional Grammar.

Francis Cornish currently teaches Syntax and General Linguistics at the Department of Language Sciences of the University of Toulouse-Le Mirail.

